Measurement

Measurement

- Careful deliberate observations of the real world
- To describe objects, events in terms of certain attributes
- It becomes a variable

Initial Observation

- Find something that needs explaining
  - Observe the real world
  - Read other research
- Test the concept: collect data
  - Collect data to see whether your hunch is correct
  - To do this you need to define variables
    - Anything that can be measured and can differ across entities or time.

Systematic Observation

The process should be replicable

i.e.

independent observers should also be able to observe and report whatever we, as researchers, observe and report.

Measurement

- Most variables we want to measure do not exist (are not tangible)
- They seldom have a single meaning

Measurement

All scientists measure:

- Direct observables
  - Eg gender, height, blood pressure, metal content

In social science we also measure

- Indirect observables
  - Eg characteristics as indicated by responses on questionnaires
- Constructs
  - Eg “success” or “compassion” measured by a scale

We need indicators and dimensions

- Indicator – sign of presence or absence of concept
  - Eg visiting children’s hospitals during Christmas can be an indicator of compassion
- Dimension
  - A specifiable aspect of a concept, eg “compassion for humans” or “compassion for animals”

Progress of Measurement

- Conceptualisation – determine different meanings and dimensions
- Nominal definition – Define the construct using a real world situation
- Operational definition – create/define the direct question on the questionnaire
- Measurement in the real world – As asked by interviewer or by questionnaire

Descriptive and Explanatory studies

- Need clear precise definitions in descriptive research
  - eg “failure rate of students”
    - define student first (short course vs part-time vs full-time)
    - will failure rate include repeating at least one subject during at least one year?
- Need less clear definitions in explanatory research
  - Eg “why is failure rate high?”

Operationalisation

- Be clear about the full range of variation
  - Eg political orientations range from very liberal to very conservative
- Degree of precision
  - How fine must distinctions be? Eg age intervals:
    16 – 20; 21- 25; 26 – 30; 31 – 35; etc,
    or
    16 – 25; 26 – 35; etc?
- Rethink dimensions
- Define variables and attributes
  - “Gender” is a variable; “Female” is a attribute

Nature of Measurement

Assignment of numbers, in terms of fixed rules, to individuals or objects to reflect differences between them in some or other characteristic or attribute.

Data types

Levels of Measurement

Categorical (entities are divided into distinct categories):

- Nominal variable:
  - a) There are only two categories
    - e.g. dead or alive.
  - b)There are more than two categories
    - e.g. whether someone is an omnivore, vegetarian, vegan, or fruitarian.
- Ordinal variable: The same as a nominal variable but the categories have a logical order
  - e.g. whether people got a fail, a pass, a merit or a distinction in their exam.

Levels of Measurement

Continuous (entities get a distinct score):

- Interval variable: Equal intervals on the variable represent equal differences in the property being measured
  - e.g. the difference between 6 and 8 is equivalent to the difference between 13 and 15.
- Ratio variable: The same as an interval variable, but the ratios of scores on the scale must also make sense
  - e.g. a score of 16 on an anxiety scale means that the person is, in reality, twice as anxious as someone scoring 8.

- Nominal
  - Categories
    - e.g., Male-female
  - Count
- Ordinal
  - Categories
  - Ordering implied
    - e.g., High-low
  - Count
- Interval
  - Equal intervals
  - No true 0
    - e.g., Degrees Celsius
  - Measurement
- Ratio
  - Equal intervals
  - True 0
  - Meaningful ratios
    - e.g., Height in inches

Data Levels – Categorical Data

Nominal
'campus

'campus'	'campus_code'
"Bellville"	1
"Cape Town"	2
"Mowbray"	3
"Wellington"	4
"Granger Bay"	5

Ordinal
'post level'

'post level'	'post level'
"research assistant"	1
"researcher"	2
"senior researcher"	3
"professor"	4
"director"	5

Data levels - Numerical Data

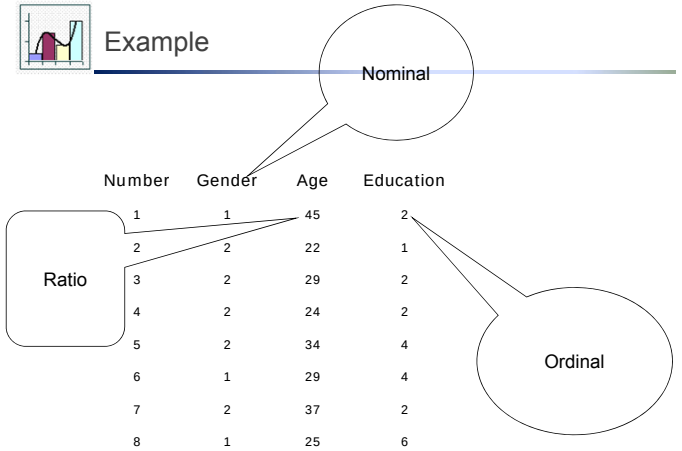
- Interval → 'temperature'

24° Celsius = 75.2° Fahrenheit

- Ratio → 'rainfall'

2.5 ml rain

Example

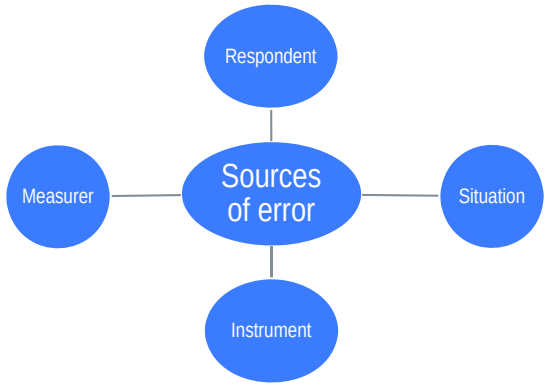


Measurement Error

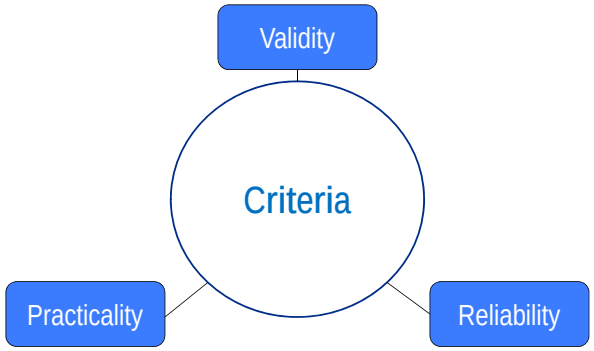
Measuring instruments measure three components:

Construct/measurement intended
Irrelevant constructs → *Systematic Sources of variation – constant for individuals*

Random Measurement Error → *Unsystematic source of variation – accidental factors - variable*



Evaluating measurement tools



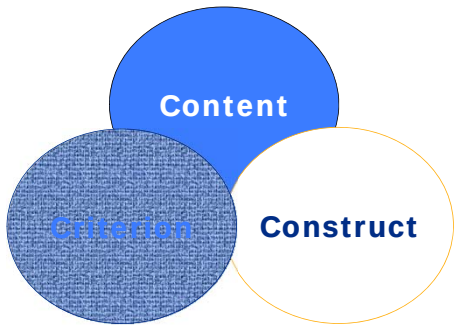
Validity

The extent to which an empirical measure adequately reflects the real meaning of the concept.

i.e.

Validity means that we measure exactly what we say we measure

Validity Determinants



Content Validity

How much a measure covers a range of meanings included within a concept, i.e.

"… the extent to which a measure represents all facets of a given social construct."

Eg a test of mathematical ability cannot only cover addition; must include multiplication, division, etc.

Construct Validity

The instrument we use to measure a variable must measure that which it is supposed to measure.

There may be more than one operational definition of the same construct.

None of the individual indicators completely succeeds in representing the construct because they also reflect other (irrelevant) constructs.

Construct Validity

Definition:

The degree to which the measuring instrument *measures the intended construct* rather than irrelevant constructs or measuring error.

Criterion-related Validity

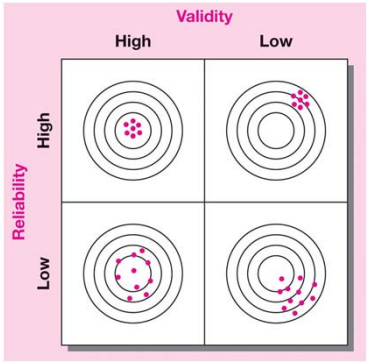
… refers to how well one variable or set of variables predicts an outcome based on information from other variables

Face Validity

… the extent to which a test is subjectively viewed as covering the concept it is supposed to measure

i.e. a test can be said to have face validity if it "looks like" it is going to measure what it is supposed to measure

Understanding reliability and validity



Reliability versus Validity

There is a tension between reliability and validity – causing a trade-off

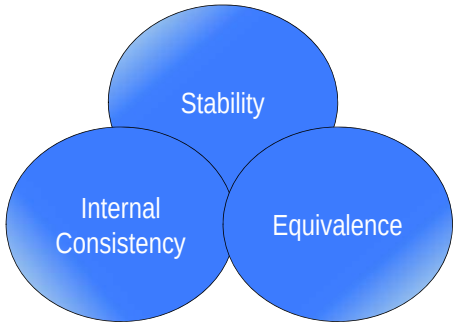
Reliability

Reliability refers to the extent to which the obtained scores may be ***generalised*** to different measuring occasions, measurement/tests forms, and measurement/tests administrators.

i.e.

Whether a particular technique, applied repeatedly to the same object, yields the same result each time.

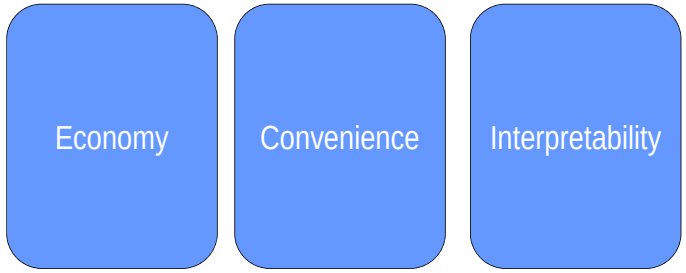
Reliability Estimates



Estimating Reliability

- Stability: Test-retest reliability – refers to the degree a measurement/test is immune to the particular measurement/test occasion. Can the results obtained be generalised to all other potential occasions.
- Equivalence: Parallel-forms reliability – is determined by interchangeable versions of a measurement/test which have been compiled to measure the same construct equally well, but with different content.
- Internal Consistency: Refers to the degree of homogeneity across the items within the measurement/test. Do the items reflect the same underlying construct
- Interrater/intercoder/tester/test or measurement reliability – refers to unreliability due to accidental inconsistent behaviour on the part of the administrator or the scorer of the test.

Practicality



Ethics of measurement

Conceptualization and measurement must never be guided by bias or preferences for particular research outcomes

Response Styles

- Validity of rating scales is negatively affected by presence of certain response styles:
  - the halo effect – Participants rated favourably because of a general favourable impression, or vice versa.
  - the severity or stringency error – All participants rated too strictly (or too leniently).
  - the error of central tendency - Raters are hesitant to assign extreme ratings and tend to place most participants in the centre of the scale.

Response Styles

- The logical error – the tendency to rate participants similarly on attributes which are incorrectly assumed to be logically related
- The proximity error – the tendency to rate attributes that appear together similarly
- The contrast error – the tendency to exaggerate the difference between themselves and the participants in respect of the attribute in question

References

- Babbie, E. 2007. *The practice of social research,* Belmont, CA, Wadsworth: Cengage Learning.
- Cooper, D. R. & Schindler, P. S. 2008. *Business Research Methods,* New York : McGraw-Hill.
- Struwig, F. W. & Stead, G. B. 2001. *Planning, designing and reporting research,* Cape Town: Pearson Education.